图书情报工作
第63卷 第8期 2019年4月

融合新兴领域知识融合过程研究
——以生物信息领域为例*

周源[1] 董放[1] 刘宇飞[1,2]
[1] 清华大学公共管理学院 北京 100084 [2] 中国工程院战略咨询中心 北京 100088

摘要：[目的/意义]技术融合是新兴产业形成和发展的驱动力，而知识融合是技术融合的前提，对知识融合过程进行研究，对于引导新兴产业形成和发展具有重要意义。[方法/过程]首先根据现有的研究构建一个采用论文引用网络表征知识融合过程的理论模型，其次根据知识融合过程每个阶段论文引用网络的特征设计验证方法，最后以融合新兴生物信息领域为例进行实证分析。[结果/结论]生物信息领域的实证分析结果表明理论模型的有效性，可为研究知识融合过程提供一种新的方法。

关键词：知识融合 知识流动 引用网络 新兴领域
分类号：G254
DOI：10.13266/j.issn.0252-3116.2019.08.015

1 引言

新兴技术是我国创新驱动发展的重要着力点，引导发展新兴技术对我国抢占未来经济和科技制高点，走上创新驱动、持续发展的轨道具有重大意义。新兴技术的形成与发展存在多种路径[1]，其中技术融合作为新兴技术产生与发展的一个新的决定性因素引发了广泛的关注和研究[2]。科学客观地分析技术融合过程对于引导新兴技术的形成和发展具有重要意义。

在F. Hacklin提出的技术融合的演化框架中，知识融合是技术融合的前提[3]，并驱动了新兴技术的形成和发展[4]。在知识融合的发展过程中，会出现不同于原有科学领域的知识外溢。这种知识外溢往往发生在现有科学领域之间，随着知识外溢程度的不断加深，新的研究范式出现，原有的科学领域便经知识融合产生新的科学领域以获得新的功能或效率改善，最终导致技术融合的出现[3-5]。如融合新兴领域生物信息的形成，生物科学领域起初多采用实验观测法进行研究，但随着观测数据的不断增多，研究者在19世纪70年代开始将信息科学的知识和方法应用于生物科学领域。随着生物科学与信息科学之间的知识外溢不断加深，并逐渐形成一种不同于生物科学和信息科学领域研究范式的一种新的研究范式，新的研究范式的出现表明了生物信息领域的产生。

现有知识融合和技术融合的研究多通过预测融合的出现和采用指标对融合现象进行截面研究。一些研究采用定性研究方法来分析融合现象，如我国的赵红州提出一种科学发现的采掘模型来解释技术演化过程出现的交叉融合[6-7]，E. Leon基于专家知识提出了一种预测融合出现方法[8]。另外一些研究则采用基于数据指标的定量研究方法，如R. Kong等通过计算两种领域之间相似度衡量领域间融合程度[9]，H. Park等也提出了基于IPC分类的各项指标识别分析领域间融合状态的算法[10]。最近一些学者开始采用聚类方法和专利引用网络相结合的方法[11]。现有的研究能够有效地从一个或多个方面识别知识融合和技术融合的状态，然而无法有效地描述和研究知识融合的具体演化过程。

论文是科学知识的载体与交流的媒介[12-13]，论文间的引用过程是在前人知识的基础上进行知识选择、进化、传播和应用的过程[14]，论文间的相互引用关系则反映了知识的流动过程[15]。论文引用网络是以论

* 本文系国家自然科学基金项目“支持技术预见的多源异构大数据融合与时序文本预测方法研究”（项目编号：91646102）和国家自然科学基金项目“面向2035的中国工程科技发展路线图绘制理论与方法研究”（项目编号：L1624045）研究成果之一。
作者简介：周源（ORCID：0000-0002-9198-6586），副教授，博士，博士生导师；董放（ORCID：0000-0003-4271-9702），博士研究生；刘宇飞（ORCID：0000-0001-9420-8811），博士后，通讯作者，E-mail：liuyufei0418@qq.com。
收稿日期：2018-06-11 修回日期：2018-12-25 本文起止页码：127-134 本文责任编辑：杜杏叶

文为节点、论文间的引用关系为连接的网络,其能够反映大数据量下复杂的知识流动,是研究知识融合过程中知识流动的有效工具。在对科学领域进行描述时, I. Sakata 指出论文引用网络是一种相比于语义分析更为有效的方法[16]。根据 Y. Kajikawa 与 M. Newman 等人的研究结果[17-18],拓扑聚类的方法可以根据论文引用网络中论文节点的聚集程度将论文引用网络划分为多个簇团,其中每个簇团均对应于一个科学领域,结合网络可视化则能够直观地展示出领域的分布状态和演化过程,是描述科学领域及其变化的有效方法。因此,相比于现有的融合状态描述的方法,本文采用引用网络可视化与拓扑聚类相结合的方法分析了知识融合的整个过程。

本文首先梳理了知识流动和领域描述方面的研究,构建了一个采用论文引用网络表征知识融合过程的理论模型;然后设计了理论模型的验证方法,并在融合新兴领域——生物信息领域进行实证分析;最终根据实证分析结果完成对理论模型的验证和完善。本文所构建的知识融合过程理论模型提供了一种新的研究知识融合过程的方法,能够为科研人员研究融合新兴领域的形成与发展过程提供一定的帮助。

2 基于论文引用网络表征的知识融合过程理论模型

2.1 知识融合理论概述

N. Rosenberg 首次提出技术融合这一概念[19],技术融合并非技术层面的单一、孤立的现象,在其产生和发展过程中,需要相关基础科学、应用科学为技术融合提供重要支持,与社会经济、法律规则等关系密切[20-21]。技术融合整个过程分为知识融合、技术融合、市场融合和产业融合四个时间阶段,其中知识融合是技术融合的前提[22]。在理想的技术融合过程中,知识融合结束技术融合随之产生并促使新的产品与市场的结合进而导致市场融合的产生以及企业间合并,最终完成产业融合[4]。

知识融合对应于技术融合存在两种形式,如图1所示,前一种形式为两种科学领域(A 和 B)融合形成新的科学领域(C),此时原有的科学领域依然存在(简称 A + B = A + B + C 型);后一种形式为两种科学领域(A 和 B)融合互补形成新的科学领域并替换掉原有的科学领域(AB)[9](简称 A + B = AB 型)。本文所构建的采用论文引用网络表征知识融合过程的理论模型只针对 A + B = A + B + C 形式的知识融合。

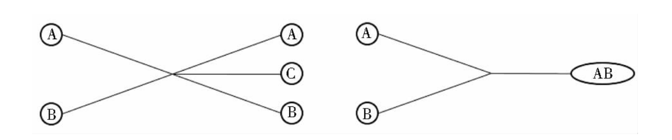


图1 两种不同的知识融合形式

2.2 知识融合过程理论模型

知识融合整个过程伴随着活跃的知识流动并会形成新的科学领域,现有的研究指出知识融合过程开始于两个不同科学领域间的知识流动,随着领域间的知识流动逐渐加深,最终当形成新的研究范式则表明融合领域的形成[3-5]。具体来说,知识融合的过程则是从现有科学领域按照各自的研究范式独立发展,到原有的科学领域间出现知识流动,最终形成新的融合领域的过程。因此,本文采用论文引用网络中的不同科学领域间相互引用表征科学领域间的知识流动,采用论文引用网络经拓扑聚类后簇团的变化表征融合过程中科学领域的变化,最终构建了一个基于论文引用网络表征的知识融合过程的理论模型,如图2所示:

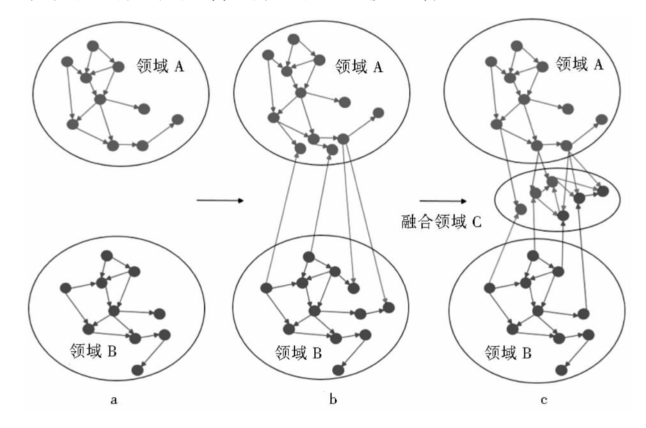


图2 知识融合过程理论模型

基于论文引用网络表征的知识融合理论模型将知识融合过程划分为以下3个阶段:

阶段1:科学领域 A、B 在阶段1均按照各自的研究范式发展,其中的知识流动为科学领域 A、B 各自领域内的知识流动,此时论文引用网络中的相互引用关系则主要为各自领域内论文间的相互引用,论文引用网络经拓扑聚类划分后的簇团也仅存在领域 A 所对应的簇团和领域 B 所对应的簇团。阶段1见图2。

阶段2:科学领域 A、B 在阶段2研究范式间出现交叉渗透,其中科学领域 A、B 开始出现领域间的知识流动,但在该阶段并未形成新的融合领域,论文引用网络经拓扑聚类划分后的簇团也仅存在领域 A 所对应的簇团和领域 B 所对应的簇团,并未出现融合领域所对应的簇团。阶段2见图2。

阶段3:科学领域A、B在阶段3经知识融合新的研究范式开始形成,标志着融合领域的出现,此时论文引用网络经拓扑聚类划分后的簇团除了领域A、B所对应的簇团之外,融合领域所对应的新簇团开始出现,本文中称融合领域对应的簇团为融合簇团。阶段3见图2。

3 知识融合过程理论模型验证方法

3.1 验证方法总体框架

知识融合过程理论模型验证方法是基于知识融合过程理论模型建立的。"知识融合过程理论模型是否符合实际知识融合过程","知识融合过程的理论模型是否完善?"这两个问题是该验证方法所要解决的问题。鉴于此,本文建立了一个知识融合过程验证方法的总体框架(见图3)。

知识融合过程理论模型验证方法总体框架分为以下3个步骤:

步骤1:论文引用网络构建。首先制定所研究领域的检索式,然后依据检索式从论文数据库中检索出相应的论文数据集,之后提取论文数据集中论文间的直接引用信息,最终完成论文引用网络的构建。

步骤2:论文引用网络分析。分别采用可视化方法识别跨领域引用,拓扑聚类方法识别融合簇团,LDA主题模型提取簇团主题词,从跨领域引用,融合簇团,簇团主题词3个方面分析所建立论文引用网络的特征。

步骤3:知识融合状态识别。根据论文数据集时间区间中每年的论文引用网络在跨领域引用,融合簇团,簇团主题词3个方面的特征,首先划分知识融合过程的阶段,之后识别融合簇团的研究主题,最终完成知识融合过程理论模型的验证和完善。

3.2 论文引用网络构建

在本文中,选取汤森路透的Web of Science(WOS)论文数据库作为论文数据源,数据库具体选择WOS中的SCI-EXPANDED,SSCI,A & HCI数据库,这三个数据库包含数千种学术期刊且引用信息完备,是进行引用信息提取最有效的数据库[23]。在进行论文数据检索时,分别选取科学领域A、B的若干关键词来表示该科学领域,根据每个领域的关键词制定检索式在WOS中检索,从而获得科学领域A、B的论文数据集。为了准确的分析科学领域A、B知识融合的过程,需确保科学领域A、B检索的数据具有相同的终止时间,且重合的时间跨度尽可能的长。由于直接引用关系更能反映知识流动和论文的聚集特征[24],因此采用Java语言解析科学领域A、B全部论文数据集中论文间的直接引用信息,以论文作为节点,直接引用关系为连接构建论文引用网络。

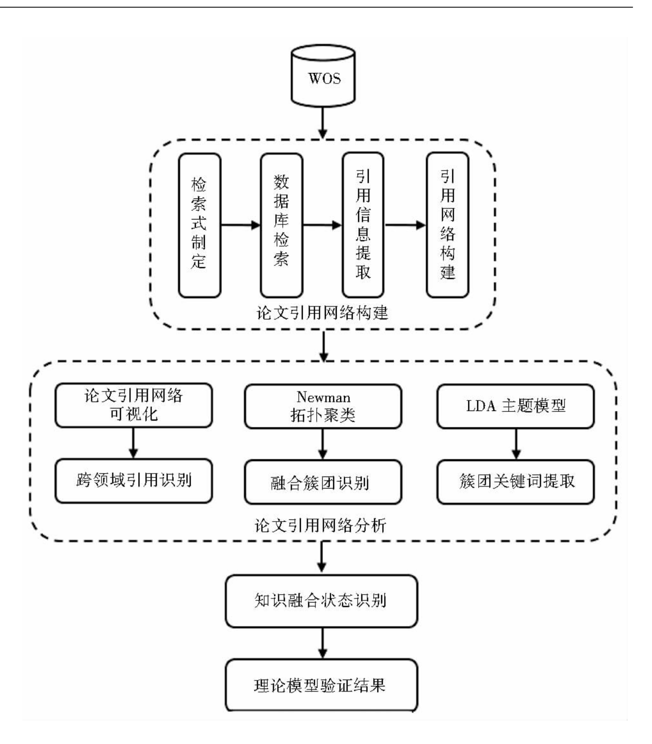


图3 验证方法总体框架

3.3 论文引用网络分析

根据知识融合过程的理论模型,知识融合过程在不同的阶段分别在论文引用网络中的跨领域引用数量,是否存在融合簇团以及簇团主题词3个方面表现出不同的特征,因此分析论文引用网络中这3个方面的特征是知识融合状态识别的前提同时也是验证方法的关键。

3.3.1 跨领域引用识别 跨领域引用识别是对论文引用网络中分别属于不同科学领域论文间的相互引用关系进行识别。不同科学领域论文间的相互引用反映了知识融合过程中领域间的知识流动,在本文中根据论文引用网络的可视化结果来识别论文间跨领域引用的数量。在可视化显示论文引用网络时,通过不同的颜色来区分引用网络中节点和连接的类别,其中节点根据论文所属科学领域的不同采用两种不同的颜色显示;连接的颜色根据连接两端论文节点的类别确定,当两个节点同属一个科学领域时节点间连接的颜色与该领域论文节点的颜色相同,当两个节点所属科学领域不同时,则节点间连接的颜色显示为白色。基于上述可视化方法开发了论文引用网络可视化软件Citation-Network Data Analyzer(CDA),见图4。采用CDA软件

检测论文引用网络中是否存在白色连接,存在白色连接即论文引用网络中存在论文间跨领域引用。

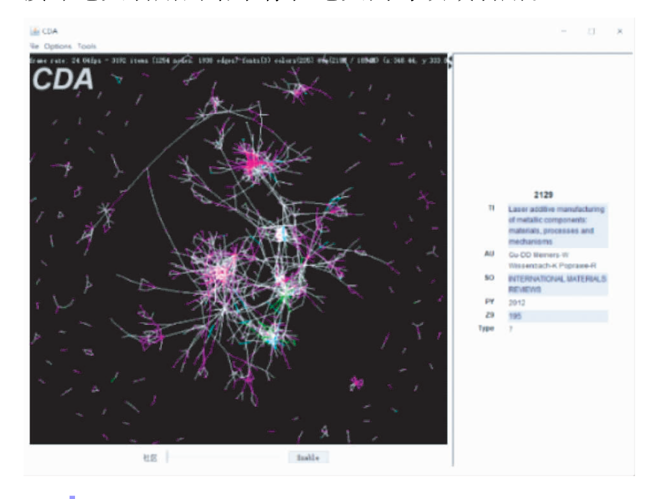


图4 CitationNetwork Data Analyzer 界面

3.3.2 融合簇团识别 融合簇团识别是通过对知识融合形成的新簇团进行识别,从而发现潜在的新科学领域。I. Rafols 在综述跨学科研究中的文献计量方法时提出技术多样性的均匀度可以反映学科融合的程度[25]。李丫丫基于技术多样性的均匀度实证分析了融合产业生物芯片产业的产生[26],娄岩在以信息技术与电动汽车技术的融合为例综述融合的测度方法时提出了多个基于技术多样性均匀度的融合测度指标[27]。知识多样性的均匀程度更高的簇团表明其包含的论文更多地来自于不同的领域,基于技术多样性的均匀度,本文通过计算簇团中每个领域论文占比快速直观地描述该簇团知识多样性的均匀程度。融合簇团的识别首先采用拓扑聚类的方法提取出论文引用网络中的全部簇团,拓扑聚类能够根据引用网络节点间的聚集特征将网络划分为多个簇团,效果见图5。在本文中选用 Newman 聚类算法来实现论文引用网络的拓扑聚类过程,该算法相比于其他聚类算法无需指定聚类簇的数量,而是按照节点间的聚集特征自动将引用网络划分为最优数量的簇团[28]。在提取出全部的簇团之后,分析每个簇团中所包含两个科学领域论文的比例,其中均匀包含两个科学领域的簇团即为融合簇团。

3.3.3 簇团主题词提取 融合簇团对应于知识融合所形成的新的融合领域,为了确定该科学领域的研究主题,采用主题词提取的方法提取融合簇团中包含论文的主题词,并通过主题词来描述该科学领域的研究主题。在本文中采用 LDA 主题模型来实现簇团主题词提取的过程,LDA 主题模型是自然语言处理中一种常用的非监督主题提取方法,其基本思想是描述文档的主题概率分布进而描述单词的主题概率分布。本文采用 LDA 主题模型提取主题词时,不考虑主题词在不同主题下的分布,因此选择主题数为 $K=1$ 的 LDA 主题模型对论文题目进行主题词提取,LDA 主题模型中的先验超参数 α 和 β 选取经验值[29] $\alpha = 50/K$, $\beta = 0.01$(即 $\alpha = 50$, $\beta = 0.01$,迭代次数 1 000 次)。

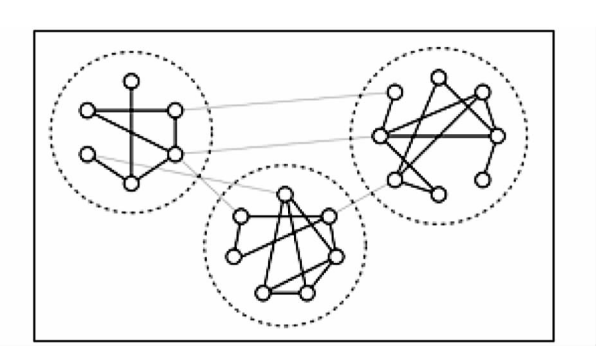


图5 拓扑聚类簇团划分

3.4 知识融合状态识别

本文所构建的基于论文引用网络表征的知识融合过程理论模型将知识融合过程划分为3个阶段。知识融合状态识别即根据实证案例某一年份论文引用网络中是否存在跨领域引用、是否存在融合簇团、簇团主题词3个方面的特征确定该年份所处知识融合的阶段。表1为知识融合过程各个阶段在论文引用网络中是否存在跨领域引用,是否存在融合簇团、融合簇团主题词3个方面的特征。

表1 知识融合过程各阶段融合状态识别

阶段	状态特征
阶段1	原有领域 A、B 间不存在论文的跨领域引用;簇团均为原有科学领域对应的簇团
阶段2	原有领域 A、B 间存在论文的跨领域引用;簇团均为原有科学领域对应的簇团
阶段3	原有领域 A、B 间存在论文的跨领域引用;簇团除了原有科学领域对应的簇团之外存在融合簇团,融合簇团论文主题词属于融合领域

通过识别分析实证案例每年知识融合的状态特征并结合阶段划分结果对所构建的理论模型进行验证。当实证案例按照其每年论文引用网络的特征可以被划分为理论模型中的3个阶段且在每个阶段符合上述理论模型中各阶段的状态特征时说明本文所构建的知识融合过程理论模型符合实际的知识融合过程。由于本文所构建的理论模型是根据现有的理论推导而来,实际的知识融合过程要比理论模型复杂很多,因此根据实证案例在知识融合过程各阶段所观察到其它特征对提出的理论模型进行完善。

4 知识融合过程理论模型实证分析

生物科学与信息科学是21世纪科学技术发展的制高点，二者经过知识融合形成的生物信息领域具有战略性新兴产业科学技术的显著特征，是一种公认的融合新兴领域。因此采用生物科学与信息科学经知识融合形成生物信息领域这一知识融合过程作为实证案例来实证分析本文所构建的采用论文引用网络表征的知识融合过程理论模型是最为合适的。

4.1 论文引用网络构建

按照方法框架的步骤1，由若干名相关领域研究人员确认生物科学和信息科学领域的关键词，检索年限设定为1995－2016年，由于生物信息领域最先在美国产生，为了降低数据噪音将检索论文的范围限定在美国。所制定的检索式为“TS＝((electrical computing) or (information technology systems) or (engineering electrical electronic) or (engineering industrial) or (software engineering) or (computer artificial intelligence) or (telecommunications) or (computer hardware architecture) or (information technology)) and CU＝USA”和“TS＝((biochemistry molecular biology) or (pharmacology pharmacy) or (biochemical research methods) or (genetics biology) or (biochemistry molecular biology) or (biotechnology) or (chemistry medicinal) or (microbiology)) and CU＝USA”。在WOS中的SCI-EXPANDED，SSCI，A&HCI论文数据库根据上述检索条件分别进行论文检索，检索结果包含生物领域论文24 319篇，信息领域论文52 254篇。提取两个科学领域全部论文数据集中论文间的直接引用信息，分别构建每年的论文引用网络时间序列。

4.2 实证分析结果

根据理论模型在知识融合不同阶段的状态特征，采用所提出的验证方法对生物科学与信息科学经知识融合形成生物信息领域的过程进行分析并与理论模型进行匹配，完成知识融合过程理论模型的实证分析过程。

在论文引用网络可视化结果中，生物科学的论文节点显示为绿色，信息科学的论文节点显示为红色。采用CDA软件对每年的论文引用网络拓扑聚类并进行可视化，根据可视化结果，生物科学与信息科学领域在1998年之前论文间的引用均为领域内的引用，不存在跨领域引用，拓扑聚类划分的簇团均为生物科学和信息科学领域各自对应的簇团，说明在1998年之前生物科学与信息科学领域的知识流动为各自领域内的知识流动。根据可视化结果，在1998年开始出现少量的跨领域引用，但经融合簇团识别在1998年未出现融合簇团，则说明到了1998年开始出现领域间的知识流动，但此时尚未融合形成新的领域，1997年与1998年的论文引用网络可视化结果见图6。根据论文引用网络在知识融合过程不同阶段的状态特征，生物科学与信息科学领域间的知识融合过程在1998年从知识融合过程的的第1阶段进入到第2阶段。

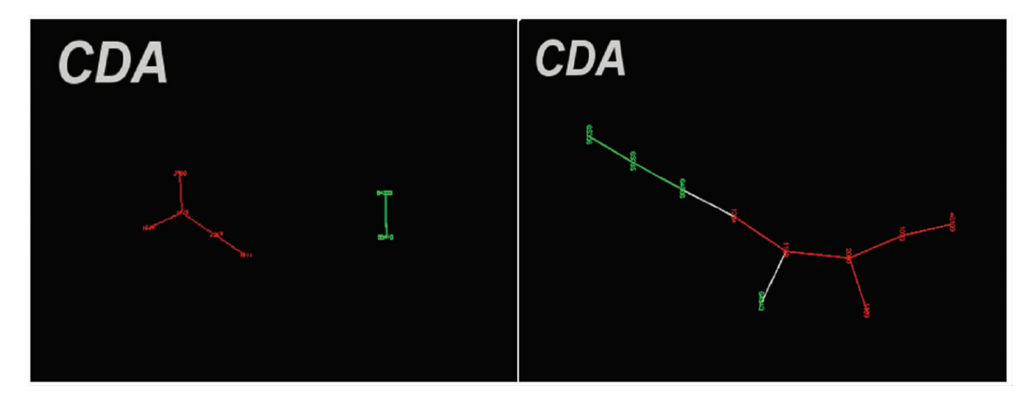


图6 1997年与1998年论文引用网络可视化结果

采用CDA软件对1998年之后的论文引用网络拓扑聚类并进行可视化，根据可视化结果，生物科学与信息科学领域在1998年与2003年之间，论文间跨领域引用的数量不断增多，2003年之前拓扑聚类划分的簇团均为生物科学和信息科学领域各自对应的簇团，并未出现融合簇团，说明1998年与2003年之间领域间的知识流动程度不断加深，但尚未融合形成新的领域；在2003年，论文间跨领域引用的数量达到一定数量，经融合簇团识别结果，论文引用网络中除了生物科学和信息科学领域各自对应的簇团，开始出现一个包含

两种不同科学领域的论文，即融合簇团，则说明到了 2003 年生物科学和信息科学经知识融合开始出现新的研究范式并形成新的融合领域。1999 年与 2003 年的论文引用网络可视化结果见图 7，2003 年出现的融合簇团在可视化结果中圈出。在 2003 年共识别出 4 个簇团，融合簇团中生物科学领域论文占比 61.3%，其他三个簇团生物科学领域占比分别为 99.9%、0.9% 和 1.8%，融合簇团包含两种领域论文的均匀程度远大于其他簇团。

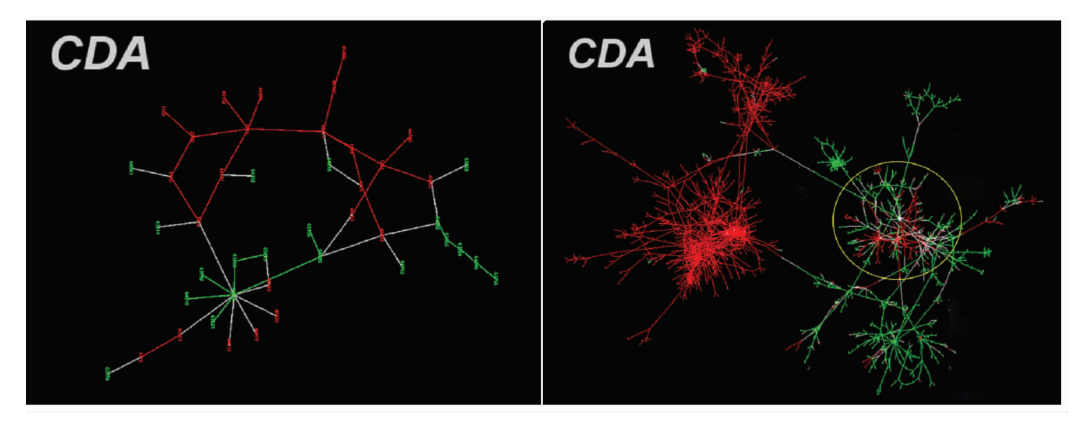


图 7　1999 年与 2003 年论文引用网络可视化结果

为了确定融合簇团所对应的科学领域，首先采用拓扑聚类识别 2003 年之后每年论文引用网络中的融合簇团，根据识别结果，从 2003 年至 2016 年每年的论文引用网络中均识别到 1 个融合簇团，2008 年与 2016 年的论文引用网络可视化结果见图 8，识别出的融合簇团在图 8 中圈出；之后采用 LDA 主题模型提取每年融合簇团中论文的主题词，提取出的主题词见表 2；最后结合生物信息领域的主题词分析生物信息领域从形成以来的演化过程。生物信息学的主要研究课题有基因组学、蛋白质组学、生物芯片等。根据生物信息领域的文献综述[30-32]，生物信息学的兴起始于 20 世纪 70 年代，从 20 世纪 70 年代初到 80 年代出现了许多生物信息学分析方法和生物信息学数据库。国际人类基因组计划（HGP）于 1990 年启动，在生命科学领域被称为阿波罗登月着陆计划，1995 年完成了第一次细菌基因组测序，2000 年 6 月 24 日世界各国参与 HGP 的 6 个国家研究机构同时宣布了一份工作草案的完成，与此同时，生物信息学在 HGP 的推动下迅速发展。结合表 2 中主题词的识别结果，在 2003 年出现基因芯片相关的主题词（gene，expression，microarray 等）；在 2004 年出现生物信息数据库相关的主题词（database，gene expression protein data 等）；在 2008 年之后则开始出现人类基因组测序相关的主题词（gene human database sequencing 等）。主题词的演化过程符合生物信息领域的形成和演化过程，说明融合簇团对应的则为融合形成的生物信息领域。根据论文引用网络在知识融合过程不同阶段的状态特征，生物科学与信息科学领域间的知识融合过程在 2003 年从知识融合过程的的第 2 阶段进入到第 3 阶段。

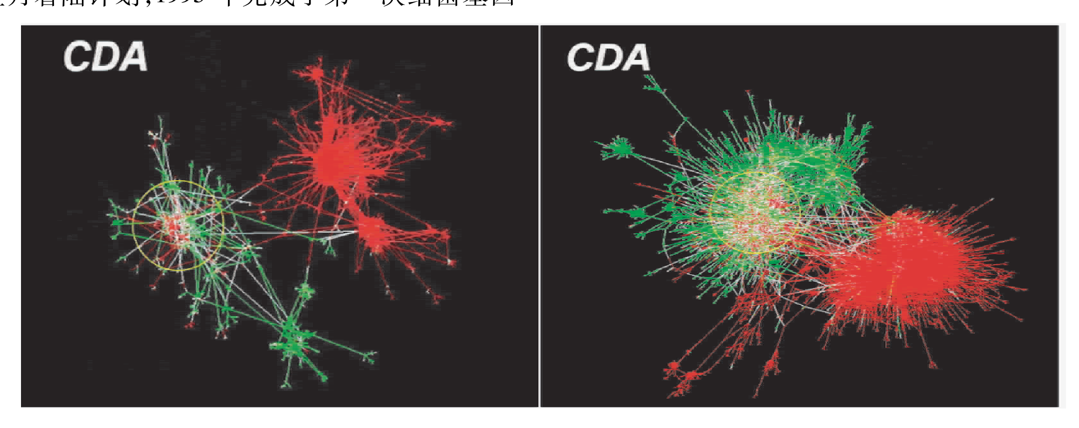


图 8　2008 年与 2016 年论文引用网络可视化结果

表 2　融合簇团论文题目主题词

年份	融合簇团主题词
2003	gene expression analysis profiling microarray discovery data protein
2004	database drug gene expression protein data information microarray
2008	gene human database sequencing biochemistry analysis microarray protein
2012	gene human analysis sequencing information database proteomics data
2016	sequencing data analysis genetic human database applications expression

根据上述分析结果,生物科学与信息科学的知识融合过程可以被划分为知识融合理论模型中的 3 个阶段,其中 1995 年至 1998 年处于知识融合理论模型的第 1 阶段,1998 年至 2003 年处于知识融合理论模型的第 2 阶段,2003 年至今处于知识融合理论模型的第 3 阶段,同时从 2003 年开始所提取融合簇团的主题词与生物信息领域相对应,并且主题词随年份的变化反映出生物信息领域的形成和演化过程。根据实证分析结果可以得出生物科学与信息科学经知识融合形成生物信息领域这一知识融合过程符合所构建的采用论文引用网络表征知识融合过程的理论模型的描述。

此外,根据实证分析结果,生物科学与信息科学经知识融合形成生物信息领域这一知识融合过程在理论模型的第 2 阶段历经 6 年时间,即从发生领域间的知识流动到形成新的科学领域所经历的时间较长。在原有理论模型中未分析融合领域形成经历的时间,实证分析结果中的这一发现从知识融合过程各阶段所经历的时间方面对理论模型进行了完善。

5　结论

本文从知识流动和领域描述的角度构建了一个采用论文引用网络表征知识融合过程的理论模型,并采用论文引用网络可视化与拓扑聚类相结合的验证方法在融合新兴领域生物信息对构建的理论模型进行验证。主要得出以下结论:①根据生物信息领域的实证分析结果证明了本文构建的采用论文引用网络表征的知识融合过程理论模型的有效性;②知识融合过程根据论文引用网络中的表征可以被划分为三个阶段,且在每个阶段具有不同的特征;③采用论文引用网络可视化与拓扑聚类相结合的方法能够有效的描述知识融合的过程。需要说明的是,本文主要从科学推动的角度描述新兴科学领域知识融合的过程,但是除此之外,一个领域的形成与发展必然还会受到经济发展、市场需求、国家政策等因素的影响,这些因素如何影响融合型新兴科学领域的形成并不在本文的讨论范围内。

参考文献:

[1] LEE S, SEOL H, PARK Y. Using patent information for designing new product and technology: keyword based technology roadmapping [J]. R&D management, 2008(38): 169－188.

[2] KIM E, CHO Y, KIM W. Dynamic patterns of technological convergence in printed electronics technologies: patent citation network [J]. Scientometrics, 2014, 98(2):975－998.

[3] HACKLIN F. Management of convergence in innovation[M]. Heidelberg: Physica-Verlag HD, 2008.

[4] CURRAN C, LEKER J. Patent indicators for monitoring convergence - examples from NFF and ICT[J]. Technological forecasting & social change, 2011, 78(2):256－273.

[5] CHO Y, KIM E, KIM W. Strategy transformation under technological convergence: evidence from the printed electronics industry [J]. Social science electronic publishing, 2015, 674(67):106－131.

[6] 赵红州. 论科学发现的采掘模型(上)[J]. 科学学与科学技术管理, 1981(2):3－5.

[7] 赵红州. 论科学发现的采掘模型(下)[J]. 科学学与科学技术管理, 1981(3):34－38.

[8] E LEON, GUILD P. Using expert knowledge to envision future converging technologies [C]//Management of engineering and technology. Portland: Portland International Center for IEEE, 2007:878－882.

[9] KONG R. Patterns and processes of contemporary technology fusion: the case of intelligent robots[J]. Asian journal of technology innovation, 2007, 15(2):45－65.

[10] PARK H, YOON J. Assessing coreness and intermediarity of technology sectors using patent co-classification analysis: the case of korean national R&D[J]. Scientometrics, 2014, 98(2):853－890.

[11] KIM D, LEE H, KWAK J. Standards as a driving force that influences emerging technological trajectories in the converging world of the internet and things: an investigation of the M2M/IoT patent network[J]. Research policy, 2017, 46(7):1234－1254.

[12] 刘先涛,朱林,何沙,等. 知识流动的综述研究[EB/OL]. 北京:中国科技论文在线 [2018－06－11]. http://www.paper.edu.cn/releasepaper/content/201406－161.

[13] 周秋菊, 杨立英, 岳婷,等. 基于期刊同被引和互引网络的学科结构和知识流动研究[J]. 情报杂志, 2014, 33(8):84－91.

[14] 梁永霞, 刘则渊, 杨中楷. 引文分析学的知识流动理论探析[J]. 科学学研究, 2010, 28(5):668－674.

[15] 王亮, 张庆普. 基于引文网络的知识流动过程与机制研究[J]. 哈尔滨工业大学学报(社会科学版), 2014(1):110－116.

[16] SAKATA I, SASAKI H, AKIYAMA M, et al. Bibliometric analysis of service innovation research: identifying knowledge domain and global network of knowledge[J]. Technological forecasting &

social change, 2013, 80(6):1085 – 1093.
[17] KAJIKAWA Y, OHNO J, TAKEDA Y, et al. Creating an academic landscape of sustainability science: an analysis of the citation network[J]. Sustainability science, 2007, 2(2):221.
[18] NEWMAN M. The structure and function of complex networks[J]. Siam review, 2003, 45(1/2):40 – 45.
[19] ROSENBERG N. Technological change in the machine tool industry, 1840 – 1910[J]. Journal of economic history, 1963, 23(4):414 – 443.
[20] NELSON R, WINTER S. An evolutionary theory of economic change[M]. Cambridge: Belknap Press of Harvard University Press, 1982.
[21] FAGERBERG J, VERSPAGEN B. Technology-gaps, innovation-diffusion and transformation: an evolutionary interpretation[J]. Working papers, 2002, 31(8/9):1291 – 1304.
[22] CURRAN C, BRÖRING S, LEKER J. Anticipating converging industries using publicly available data[J]. Technological forecasting & social change, 2010, 77(3):385 – 395.
[23] ITTIPANUVAT V, FUJITA K, SAKATA I, et al. Finding linkage between technology and social issue: a literature based discovery approach[J]. Journal of engineering & technology management, 2014, 32(2):160 – 184.
[24] IWAMI S, MORI J, KAJIKAWA Y, et al. Comparison of indicators to detect emerging researches using time transition in quasicrystals[C]//IEEE international conference on industrial engineering and engineering management. Beijing: IEEE, 2014:48 – 52.
[25] RAFOLS I, MEYER M. Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience[J]. Scientometrics, 2009, 82(2):263 – 287.
[26] 李丫丫, 赵玉林. 基于专利的技术融合分析方法及其应用[J]. 科学学研究, 2016, 34(2):203 – 211.
[27] 娄岩, 杨培培, 黄鲁成, 等. 基于专利的技术融合测度方法——以信息技术与电动汽车技术的融合为例[J]. 现代情报, 2017(8):144 – 155.
[28] NEWMAN M. Modularity and community structure in networks[C]//APS march meeting. Baltimore: American Physical Society, 2006:8577 – 8582.
[29] GRIFFITHS T, STEYVERS M. Finding scientific topics[J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101: 5228 – 5235.
[30] HOGEWEG P. The roots of bioinformatics in theoretical biology[J]. Plos computational biology, 2011, 7(3):e1002021.
[31] GRAU J, BENGAL I, POSCH S, et al. VOMBAT: prediction of transcription factor binding sites using variable order Bayesian trees[J]. Nucleic acids research, 2006, 34(Web Server issue):W529 – W533.
[32] FLEISCHMANN R, ADAMS M, WHITE O, et al. Whole-genome random sequencing and assembly of haemophilus influenzae rd.[J]. Science, 1995, 269(5223):496 – 512.

作者贡献说明:

周源:负责框架设计,论文修改及撰写指导;
董放:负责实验设计与论文撰写;
刘宇飞:负责观点提炼,论文修改。

Research on the Process of Knowledge Fusion in Emerging Fields: A Case Study of Bioinformatics

Yuan Zhou[1] Fang Dong[1] Yufei Liu[1,2]
[1] Tsinghua University, School of Public Policy and Management, Beijing 100084
[2] Chinese Academy of Engineering, The CAE Center for Strategic Studies, Beijing 100088

Abstract: [**Purpose/significance**] Technology fusion is the driving force of new industries' formation and development. While the knowledge fusion is the prerequisite of technology convergence, it is of great significance in guiding the formation and development of new industries. [**Method/process**] Firstly, this paper built a theoretical model which using the citation network to characterize the knowledge fusion process based on the existing research. Then, based on the characteristics of paper citation network in each stage of knowledge fusion, the paper introduced a verification method. Finally, it conducted an empirical analysis in the field of bio-information technology. [**Result/conclusion**] The results of empirical analysis in the field of biological information show the validity of the theoretical model. Therefore, It can be deemed as a new method on studying the knowledge fusion.

Keywords: knowledge fusion knowledge flow citation network emerging field